



Assessments of Composite and Discrete Sampling Approaches for Water Quality Monitoring

Rachel Cassidy¹ • Phil Jordan^{2,3} •
Marianne Bechmann⁴ • Brian Kronvang⁵ •
Katarina Kyllmar⁶ • Mairead Shore^{3,7}

Received: 19 February 2017 / Accepted: 26 March 2018 /

Published online: 12 April 2018

© The Author(s) 2018

Abstract Achieving an operational compromise between spatial coverage and temporal resolution in national scale river water quality monitoring is a major challenge for regulatory authorities, particularly where chemical concentrations are hydrologically dependent. The efficacy of flow-weighted composite sampling (FWCS) approaches for total phosphorus (TP) sampling ($n = 26$ –52 analysed samples per year), previously applied in monitoring programmes in Norway, Sweden and Denmark, and which account for low to high flow discharges, was assessed by repeated simulated sampling on high resolution TP data. These data were collected in three research catchments in Ireland over the period 2010–13 covering a base-flow index range of 0.38 to 0.69. Comparisons of load estimates were also made with

Highlights

- River flow-weighted composite and discrete sampling approaches were assessed for chemical load
- The assessments were made on near continuous total phosphorus and discharge data
- Data were collected in three rivers of varying hydrological conditions over 3 years
- The accuracy of the composite methods were dependent on the flow constants used
- The methods show where trade-offs could be made to optimise load calculations with small datasets

✉ Rachel Cassidy
rachel.cassidy@afbini.gov.uk

¹ Agri-Environment Branch, Agri-Food and Biosciences Institute (AFBI), Belfast BT9 5PX, Northern Ireland

² School of Geography and Environmental Sciences, Ulster University, Coleraine BT52 1SA, Northern Ireland

³ Agricultural Catchments Programme, Teagasc, Johnstown Castle, Wexford, Ireland

⁴ NIBIO, Norwegian Institute of Bioeconomy Research, Postbox 115, N-1431 Ås, Norway

⁵ Department of Bioscience, Aarhus University, 8600 Silkeborg, Denmark

⁶ Department of Soil and Environment, Swedish University of Agricultural Sciences, 750 07 Uppsala, Sweden

⁷ Wexford County Council, Carricklawn, Wexford, Ireland

discrete (set time interval) daily and sub-daily sampling approaches ($n=365$ to >1200 analysed samples per year). For all years and all sites a proxy of the Norwegian sampling approach, which is based on re-forecasting discharge for each 2-week deployment, proved most stable (median TP load estimates of 87–98%). Danish and Swedish approaches, using long-term flow records to set a flow constant, were only slightly less effective (median load estimates of 64–102% and 80–96%, respectively). Though TP load estimates over repeated iterations were more accurate using the discrete approaches, particularly the 24/7 approach (one sample every 7 h in a 24 bottle sampler - median % load estimates of 93–100%), composite load estimates were more stable, due to the integration of multiple small samples ($n=100$ –588) over a deployment.

Keywords Sampling · Catchments · Phosphorus loads

1 Introduction

The uncertainty of determining the chemical concentration characteristics of dynamic river conditions over lengthy timescales has been discussed over several decades (Brewin et al. 1996; Minaudo et al. 2017; Rode and Suhr 2007; Stevens and Smith 1978). This is largely due to the changing nature of flow regimes, linked to how chemicals are transported, concentrated and diluted by hydrological flow paths and instream processes (Jarvie et al. 2012; Marcé et al. 2018; Neal et al. 2012; Nimick et al. 2011). Very long time series of buffered chemical parameters (i.e. unlikely to excessively change with changing flow regime), especially in catchments with high Base-flow Indices (BFI), are more likely to provide a stable characteristic (Kronvang and Bruhn 1996).

Chemical concentration characteristics are useful to conceptualize the status or health of river sites, and can be used with ecological metrics to infer cause and effect from, for example, background physiographic conditions and/or upstream pollution potential (Melland et al. 2018; Pellerin et al. 2016). The product of chemical concentration and river discharge can also be used to infer upstream to downstream chemical loads in expressions of mass per unit time. This load expression has similarly been the subject of much research in recent decades and especially related to sediment and nutrient transport from land to water (Quilbé et al. 2006; Worrall et al. 2012a, b). More recently, the nutrient output load from catchments has been considered an important measurement to gauge the efficacy of catchment nutrient input load reduction measurements under national and international legislation to protect water resources from eutrophication (Bechmann and Øgaard 2013; Stamm et al. 2014; Vervloet et al. 2018).

‘True’ chemical output load requires a summation of instantaneous and synchronous chemical concentration and river discharge measurements over the time scale being investigated. While river discharge can be measured on a near continuous basis (with data points interpolated between sub-hourly resolutions), synchronization with chemical concentration is more difficult. The essential issue controlling uncertainty in load is the hydrological dependency of individual chemical parameters during different flow regimes. As load is multiplicative, a chemical that increases in concentration with increases in discharge (such as non-point phosphorus (P)) will result in a much higher (than the starting point) load. Chemicals that show no or little change in

concentration with increasing discharge, or a decrease in concentration with decreasing discharge (such as nitrogen (N)) will indicate little change or a stable load with time.

The use of interpolation and extrapolation models for dealing with discharge and chemical asynchronicity has been reported and assessed for P loads (Audet et al. 2014; Cassidy and Jordan 2011; Skeffington et al. 2015; Worrall et al. 2012a, b). These models deal with lower resolution chemical monitoring alongside near continuous discharge measurements and results show unacceptable uncertainty for load estimation when those chemical concentration data are reduced beyond at least daily and often sub-daily resolutions depending on catchment size and hydrological characteristics.

Semi-autonomous analyzers and probes can be used to fully synchronize river discharge and chemical concentration data for 'true' load estimation but are often only limited to specific research programmes due to initial costs and on-going maintenance. Additionally, national and international scale water quality monitoring programmes are designed to be as spatially representative as possible. However, there is a risk that uncertain load estimates based on sparse data will lead to erroneous reviews of environmental policy (Harris and Heathwaite 2005), in situations where load estimates are important, by not considering the time element (i.e. a space for time compromise). Load specific monitoring programmes, however, can thread a line between resource allocation in terms of space and the need to consider time with regard to issues such as hydrological dependency and synchronicity. Load specific sampling programmes are designed to note changes in catchment chemical exports over set time-scales and are often collated to annual exports. The data can then be appropriately compared with input load data and assessed against targets. For lakes, transitional and coastal receiving waters accurate assessment of annual river P loads is also essential in understanding and assessing internal cycling and chemical interactions.

The flow-weighted composite sampling (FWCS) method is load specific, using discharge data in real time to regulate the automated subsampling of water after the passage of a certain volume of water, i.e. using a flow-constant. Sub-samples are composited and this overall sample is analysed as a flow-weighted mean concentration which, when multiplied by the total discharge volume in the period provides the period load, L_p .

$$L_p = Q_{tot} \cdot \frac{\sum_{i=1}^n C_i}{n} \quad (1)$$

where C_i is the i^{th} sample, taken after a flow constant (defined as k_q and determined in Section 2) has passed, up to a total of n samples over the deployment and Q_{tot} is the integrated discharge over the deployment period from t_{start} to t_{end} .

$$Q_{tot} = \int_{t_{\text{start}}}^{t_{\text{end}}} Q_t dt \quad (2)$$

Sampling periods of days to weeks can be employed and the integration of each can, for example, provide the overall annual load estimate. In some cases a time-weighted composite sampling (TWCS) method may be applicable, either alone (during periods of stable flow) or in combination with a FWCS approach. In systems with a large range in annual discharge, a FWCS approach may under-represent fluxes during low flow periods due to long durations between flow-based samples. Under these conditions a TWCS method may be more appropriate, following Eqs. 1 and 2 but

with C_i as the i^{th} sample taken after a constant time interval has elapsed (defined as k_t and determined in Section 2). Alternatively, a TWCS approach used in isolation in systems with a large range of low to high flows is likely to bias load estimates to the lower flows and under-estimate annual chemical load.

The trade-off with composite sampling is that including a definition of hydrological change compromises parameter concentration characteristics on instantaneous scales. A further trade-off is that only conservative parameters can be monitored, such as total nutrients (i.e. TP and total N) and total sediments, as exchange in soluble and particulate species is likely in a composite sampling bottle, even when refrigeration is possible. Nevertheless, the export of total compositions of nutrients and sediments, for example, is an important consideration when developing catchment strategies to minimize or control these losses from land (Bechmann et al. 2008).

Establishment of the flow constant (k_q) is a key step in programming a FWCS scheme and also establishing the TWCS time constant (k_t) for low flows, where appropriate. Both are based on an analysis of long-term discharge or hydro-meteorological records. In this study, several TP FWCS procedures (some also incorporating a TWCS step) were assessed that have been part of periodic or long term monitoring programmes in Scandinavia. The assessment, principally on the flow-constant used, was made on 3 years of extant high-resolution (20 min) synchronous discharge and TP concentration data from three agricultural catchments in Ireland (Fealy et al. 2010). To account for the trade-off between loss of instantaneous concentration characteristics and inclusion of hydrological change in the load estimates, further methods were used that incorporate elements of each, tested previously in a flashy catchment (Cassidy and Jordan 2011; Jordan and Cassidy 2011), but applied here in catchments with a range of BFI_s.

2 Methods

Total P and discharge data from three agricultural catchments in Ireland (Table 1) were collected over 2010–2013 using the experimental infrastructure described in (Wall et al. 2012). The three catchments, Castledockerell and Ballycanew in Co. Wexford and Timoleague in Co. Cork (Fig. 1), represented a range of hydrological and P response conditions largely controlled by ranges in soil drainage characteristics (poor to free draining) and land use (Mellander et al. 2012; Murphy et al. 2015). Base-flow indices, reflecting characteristics, ranged from 0.38 to 0.69 over the 3 years.

Total P was measured in river water on a 20 min time-step by Hach Phosphax Sigma bankside analysers using a molybdate antimony-ascorbic acid procedure following heated digestion with sulphuric acid (Eisenreich et al. 1975). Discharge was monitored at non-standard Corbett flat-v weirs rating water level records from vented pressure sensors (OTT Nautilus-mini) with area-velocity measurements (OTT ADC) in the WISKI-SKED software. Discharge from these catchments was generated by rainfall only and so processes relating to snow or snow melt, for example, are not considered. All data were held in the WISKI 7 database management system and records linked to synchronous discharge and concentration measurements.

For this sub-hourly data the total ‘true’ load (L_T) from each catchment was calculated as:

$$L_T = \int_{t_{\text{start}}}^{t_{\text{end}}} Q_t C_t dt \quad (3)$$

Table 1 Summary characteristics for the three catchments based on Fealy et al. (2010) and the discharge time series. Baseflow index estimate is based on average daily flows, using smoothing and separation rules so that $BFI = V_a/V_b$ where V_a is the recorded hydrograph and V_b is the baseflow line, following Gustard et al. (1992)

	Area (ha)	Land use & Stocking rates (in livestock units)	P loading $\text{kg ha}^{-1} \text{ yr}^{-1}$	Soil Type	Drainage	Year range	Base flow index (BFI)	Total discharge $\times 10^6 \text{ m}^3$
Ballycanew	1207	Grassland: 77% Arable: 20% Stocking rate: 1.27 ha^{-1}	Organic: 14.848 Human: 0.155	Undifferentiated gley (71%); Typical brown earth (16%)	Poorly draining	2010–11 2011–12 2012–13 2010–13	0.39 0.38 0.39 0.38	3.78 5.81 5.52 15.11
Castledockerell	1116	Grassland: 39% Arable: 54% Stocking rate: 1.02 ha^{-1}	Organic: 9.710 Human: 0.178	Typical Brown earth (72%); Shallow Brown earth (14%)	Free draining	2010–11 2011–12 2012–13 2010–13	0.64 0.64 0.60 0.63	4.12 5.80 5.29 15.21
Timoleague	758	Grassland: 84% Arable: 6% Stocking rate: 2.33 ha^{-1}	Organic: 23.214 Human: 0.178	Typical Brown earth (91%); Shallow Brown earth (6%)	Free draining	2010–11 2011–12 2012–13 2010–13	0.59 0.69 0.59 0.63	3.47 5.01 5.19 13.67

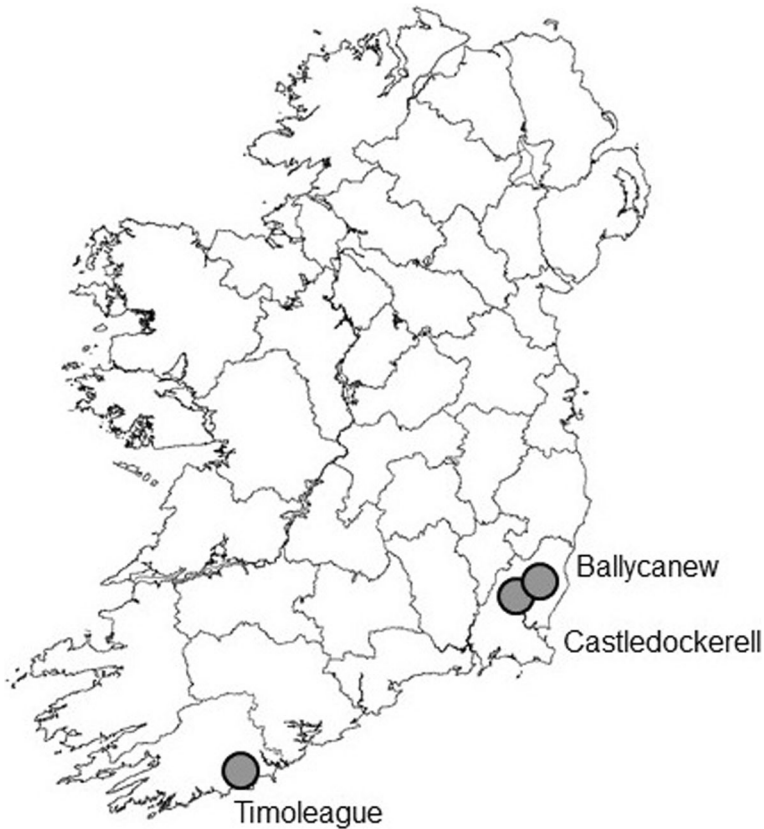


Fig. 1 Ireland showing county boundaries and locations of the three study catchments. Timoleague is located in County Cork; Castledockerell and Ballycanew are located in County Wexford

where, over a sampling period (t_{start} to t_{end}), Q_t is the instantaneous discharge and C_t the instantaneous concentration at sample time, t . This was used as the basis for evaluating all the sampling approaches trialled.

2.1 Composite Sampling and Load Estimation

The FWCS procedures based on schemes in Denmark, Norway and Sweden were assessed on the extant data. The approach was common to each scheme (Eqs. 1 and 2) but the method of assigning and managing flow and time-constants (k_q or k_t , respectively) was different. For each catchment a training set of discharge data was required to assess the likely passage of water in

Table 2 Derivation of sampling flow constants for the Swedish Monitoring Programme (based on 3 year discharge records) for the 3 catchments

		Castledockerell	Ballycanew	Timoleague
Maximum volume (m^3) in 2 week period	$Q_{\text{max}14}$	750,000	925,000	750,000
Sampling trigger volume for FWCS [m^3]	$k_q = \frac{Q_{\text{max}14}}{388}$	1276	1573	1276
Flow threshold for TWCS ($\text{m}^3 \text{s}^{-1}$)	Q_{lim}	≤ 0.0295	≤ 0.0364	≤ 0.0295

any given sampling period and each flow-constant was, therefore, catchment specific. Time-weighted composite sampling approaches, used in isolation, were not assessed in this study as all rivers were known to exhibit a high range in flows (from low to high) and previous comparisons in similar Irish rivers had demonstrated significant under-prediction compared with FWCS calculations (Macintosh et al. 2011).

2.1.1 The Swedish Approach

The Swedish approach, which was implemented on eight national monitoring catchments between 2004 and 2006 (Kyllmar et al. 2014), uses a combination of FWCS and TWCS (Eqs. 1 and 2) over a 2-week refrigerated auto-sampler deployment. During periods of low flow, sampling is time-proportional with a subsample taken every 6 h ($k_t = 6$ h); otherwise sampling is flow-proportional with a sample taken after a set flow volume has passed (k_q , given for each catchment in Table 2). In comparison to grab sampling in clay soil dominated Swedish catchments estimated losses of P were found to be higher using the FWCS approach (Kyllmar 2009), though with little change for N.

The threshold discharge to alternate from flow to time-based sampling is determined by the minimum volume of flow observed in any 2-week period for the specific river during long-term discharge monitoring. The equivalent flow rate (m^3/s) is then used as the limit, Q_{lim} , and given for each catchment in Table 2.

For FWCS in this approach the volume at which a sample is taken is apportioned from the maximum total flow in any 2 week period in the discharge time series, divided by the total number of 17 ml samples which can be stored in a 10 L collection flask ($n = 588$). A minimum of 475 ml was required for laboratory analysis so at least 28 sub-samples are necessary over the 2 week deployment period. Use of time-proportional sampling during prolonged dry periods ensured sufficient sample volume was maintained.

This was replicated in the sampling algorithm in this study by calculating the flow volume as the integral of the measured flow from the start of each 2 week sampling period (Eq. 2). The algorithm switches from FWCS to TWCS when the recorded flow is below the identified threshold ($Q_t < Q_{\text{lim}}$). As the time series is discrete the closest recorded measurement to the required flow volume (for FWCS) or 6 h timestamp (TWCS) is identified and the corresponding TP concentration stored. The flow volume is reset to zero and the computation proceeds again until the threshold volume or time interval is reached or exceeded, alternating between a flow-proportional (using flow constant k_q) or time-proportional (using time constant k_t) routine as modulated by discharge. Sampling continues until the end of the 2 week period or when the sampler capacity (588 samples) is reached. Flow volume is totalled for the 2 week period. The stored concentrations from each sampled timestamp are averaged and the load obtained from the product of the average concentration and the total flow for the 2 week period (Eqs. 1 and 2).

2.1.2 The Norwegian Approach

The Norwegian monitoring approach (Deelstra et al. 2013) uses forecasted rainfall and its relation to river flow in the calculation of the trigger volume (k_q) at which a sample is collected over a 2 week autosampler deployment. A 2 week rainfall forecast is obtained prior to set-up and used to estimate the corresponding river discharge. From this, the flow volume required to obtain a composite of 100 sub-samples over the 2 week deployment is set as a threshold upon

which a sample is taken. If the sampler capacity is reached before the end of the 2 week period the samples are collected early, the forecasting simulation repeated and a new 2 week monitoring period initiated from that date. In the Norwegian deployments, the rainfall forecasting is accurate enough so that the sampler capacity is only rarely exceeded before the end of the 2 week sampling period (Marianne Bechmann, personal communication). A phone-warning is automatically sent to the technical staff responsible for the station before the sampler capacity is reached.

As it was not feasible to replicate the weather forecasting aspect of this in the sampling algorithm, an alternative proxy, which should hold comparable degrees of uncertainty, was used instead. The approximation is to assume the forecast and resulting river discharge is accurate for the first week, so the river discharge during this period, Q_w , is calculated using Eq. 2 with t_{start} to t_{end} as the first week of discharge records and the flow constant, k_q , then calculated as

$$k_q = \frac{2 \cdot Q_w}{n_s} \quad (4)$$

where Q_w is extrapolated to a 2 week period and apportioned over the 100 possible samples, n_s , before autosampler capacity is reached.

If the sample maximum is reached before the end of the 2 week period, sampling stops, an updated 'forecast' is produced and the sub-sampling threshold re-set and the deployment restarted using the prediction for the 2 weeks from that date.

2.1.3 The Danish Approach

In Denmark, weekly flow proportional sampling was incorporated within the Danish Nationwide Monitoring Programme from 1998 until 2003 (Andersen et al. 2005). The monitoring network comprised 24 predominantly agricultural catchments in which composite samples of river discharge were acquired over a weekly deployment. An autosampler with 24 1 L bottles was deployed weekly, and programmed to abstract a sample each time a threshold discharge volume (k_q) was exceeded. Samples were retrieved every 7 days and transported to the laboratory for analysis. The autosampler was re-set with new parameters each month.

The threshold for sampling (k_q) was determined based on the median monthly flow rate, F_{med} (m^3/s), over and up to 10 years of stream discharge as

$$k_q = \frac{s \cdot F_{\text{med}}}{n_s} \quad (5)$$

where s is the number of seconds in a weekly deployment and n_s is the number of samples possible when flow is apportioned across 24×1 L bottles, each with capacity for 12 samples of 75 ml when filled to the 900 ml limit required for analysis ($n_s = 288$).

Where the weekly flow exceeds the estimate based on F_{med} then the sampler capacity will be exceeded, leaving part of the period unrecorded. In that case the mean flow-weighted concentration from the part of the week sampled is used as a proxy for the 'true' weekly mean concentration.

The numerical algorithm replicated this variant of the FWCS scheme (Eq. 1) on the time series for each catchment. Median monthly flow rates were calculated over 3-year records of flow for each catchment and from this the estimated weekly flow from which the threshold

flow volume (k_q) is determined, was calculated (Eq. 5). The flow volume, Q_{tot} (Eq. 2), calculated as the integral of the measured flow, is accumulated from the start of each sampling period. As the time series is discrete the algorithm sets a stop point once the flow exceeds or is equal to the threshold volume, and takes the concentration value corresponding to that timestamp. For the 20 min concentration time series the sampled concentration is always within 10 min of the time at which a sub-sample would have been abstracted by an autosampler in the field. The flow volume is reset to zero and the computation proceeds again until the threshold volume is reached or exceeded. Sampling continues until the end of the week or when the sampler capacity ($n_s = 288$) is reached. The total volume of flow for the entire 7 day period is determined, the stored concentrations from each sampling point over the week are averaged and the weekly load obtained from the product of the average concentration and the total flow for the 7 day period (Eqs. 1 and 2).

2.1.4 Annual Load Estimation for Composite Methods

Total annual loads for the Danish, Norwegian and Swedish approaches were calculated from the product of the total flow and the concentration from analysis of the composite samples (Eqs. 1 and 2). This was summed to give an annual estimated load

$$L_e = \sum_{i=1}^n L_{p_i} \quad (6)$$

where L_{p_i} is the i^{th} estimated load for each of n contiguous annual deployments, and compared with the true load L_T (Eq. 3).

2.2 Discrete Sampling Methods

For comparison with the Danish, Swedish and Norwegian composite sampling approaches three discrete sampling approaches were also replicated following a 7 h, 24 h and midday sampling interval during a weekly deployment scenario of a programmed autosampler. These methods had been tested previously (Jordan and Cassidy 2011) in a small flashy catchment (5km^2) and applied here in larger catchments with varying base-flow indices.

The 7-hourly sampling interval follows work in the Plynlimon experimental catchment in Wales which used an automatic water sampler with a 24 bottle capacity set to sample on a 7 h basis (Halliday et al. 2012; Neal et al. 2012). This 24/7 configuration is convenient to deploy, retrieve and analyse as part of a weekly sampling round and has been shown to provide accurate estimates of nutrient loads in headwater catchments, partially as a consequence of cycling through all times of day and night over the week and capturing something of the diurnal variations in such systems (Halliday et al. 2012; Jordan and Cassidy 2011).

Three additional variants that were evaluated include (i) a 24 h sampling programme with samples taken at any hour to return a total of 7 discrete samples over a week (ii) a 12 h sampling programme with a sample taken at any hour to return 14 discrete samples over a week and (iii) a 24 h sampling programme with a single sample taken within 1 h of midday (11:00–13:00 h). These variants allow some assessment of the sensitivity of sampling to diurnal variations in each catchment.

The paired flow and TP data were sampled every 7, 12 and 24 h and around midday from an initial starting time (which was iterated through the first week of the time series to generate all possible sample sets from the sub-hourly time series). The paired samples of flow and TP

were then used to estimate annual load, L_E , using the flux-based first-choice Paris Commission algorithm (PARCOM - Method 5 in (Littlewood et al. 1998)), where:

$$L_E = \frac{K \sum_{i=1}^n (C_i Q_i)}{\sum_{i=1}^n Q_i} \cdot Q_r \quad (7)$$

$$\text{with } Q_r = \frac{\sum_{k=1}^N Q_k}{N} \quad (8)$$

where C_i is the TP concentration (mg L^{-1}) for the i^{th} sample of a total of n possible samples according to the set time interval between samples, Q_i is the instantaneous discharge ($\text{m}^3 \text{s}^{-1}$), and Q_r is the average discharge, based on higher frequency discharge records (N measurements in total), over that sampling duration. K is a constant that accounts for the duration of the record. Q_k is the recorded discharge at 20 min intervals. The sample interval, n , is 7 h or 24 h depending on the discrete method used.

The estimated load was aggregated for each sample set across the three sampling strategies and aggregated for comparison with the true load, L_T , (Eq. 3).

3 Results and Discussion

3.1 Composite FWCS Performance

Of the three scenario composite approaches assessed the Danish monitoring strategy tended to over- or under-estimate by the most across all catchments and years (Table 3, Fig. 2). The median estimated annual loads were 98.2, 64.1 and 75.7% of the total for Ballycanew, 101.6, 86.9 and 86.7% of the total for Castledockerell and 99.5, 81.9 and 90.0% of the total (L_T) for Timoleague, over the hydrological years (October – September) 2010–11, 2011–12 and 2012–13, respectively. The best load estimates were achieved in the driest year (2010–11) and in the catchments with higher BFI (3 year mean BFI was 0.63 in Timoleague and Castledockerell compared to 0.38 in Ballycanew). Interannual rainfall variations, which can be considerable, led to a mismatch between the flow constant used for the FWCS and the weekly flow. For most weeks in a drier year the monthly threshold for sampling (based on the median monthly discharge over the 3 year record) is adequate to capture the entire week, compared to wetter years where higher flows lead to exceedance of the sampling capacity ($n_s = 288$) by mid-week, leaving the remaining days unrecorded. In Ballycanew, which is poorly drained and has a flashier response to rainfall events (BFI = 0.38) some significant diffuse transfers occurring at the end of weeks were missed entirely, leading to larger underestimates overall in this catchment, particularly in 2011–12 when the sampling capacity was exceeded before the end of the week on 51% of deployments. In 2010–11, the driest year, the Danish approach performed better than on other years where sampler capacity was sufficient for 66–74% of weekly deployments over all sample sets. In applications in Denmark the BFI of monitored sites was typically >0.7 (Brian Kronvang, personal communication).

The Swedish and Norwegian sampling approaches were more consistently accurate. Loads estimated from the Norwegian approach were between 92.7 and 97.9% accurate across the three catchments and years and appear unaffected by the variations in hydrology between the

Table 3 Median and interquartile range differences for the percentage of load estimated for the three catchments

n	sample sets	Median	25th-Median	75th-Median	Median	25th-Median	75th-Median	Median	25th-Median	75th-Median
Castledockerell 2010–11										
169	Swedish %	95.89	-0.28	0.33	91.06	-0.41	0.43	79.94	-0.20	0.36
169	Norwegian %	97.13	-0.29	0.44	92.72	-0.34	0.60	92.67	-1.59	0.73
169	Danish %	101.64	-0.79	0.86	86.85	-7.18	7.24	86.65	-7.80	5.83
171	24–7 Sampling %	99.91	-2.47	3.05	93.40	-1.20	7.88	93.91	-11.86	20.09
171	24 h sampling %	93.28	-6.71	10.68	86.48	-8.01	16.33	84.18	-14.19	10.97
171	12 h Sampling %	96.58	-3.62	9.79	92.71	-0.34	0.57	86.36	-10.37	15.59
211	Midday Sampling %	89.95	-5.42	17.83	124.80	-10.55	10.89	60.39	-1.76	2.25
Ballycanew 2010–11										
169	Swedish %	91.12	-0.13	0.42	90.34	-0.12	0.22	87.63	-0.63	1.03
169	Norwegian %	97.07	-0.61	0.42	96.35	-0.71	0.08	86.74	-0.72	1.06
169	Danish %	98.20	-1.68	1.81	64.06	-7.32	11.44	75.65	-6.46	7.62
171	24–7 Sampling %	98.14	-1.72	3.55	98.12	-0.89	0.92	100.11	-0.75	2.05
171	24 h sampling %	96.34	-5.07	8.82	98.61	-7.19	6.22	100.60	-5.74	3.93
171	12 h Sampling %	98.91	-5.27	7.79	98.59	-2.89	2.31	102.30	-6.81	2.93
211	Midday Sampling %	96.25	-2.40	1.94	107.82	-1.59	2.24	98.35	-2.01	1.95
Timoleague 2010–11										
169	Swedish %	94.93	-0.47	0.46	91.02	-0.22	0.32	90.07	-0.22	0.23
169	Norwegian %	97.93	-0.39	0.48	95.06	-0.33	0.93	96.85	-0.61	0.60
169	Danish %	99.50	-3.78	1.60	81.90	-3.60	10.86	89.97	-3.49	3.37
171	24–7 Sampling %	97.13	-1.31	8.08	96.55	-5.67	11.38	97.38	-8.04	4.61
171	24 h sampling %	96.27	-7.30	8.06	93.37	-4.22	6.31	91.24	-6.73	18.14
171	12 h Sampling %	97.51	-4.22	7.01	93.66	-3.89	5.84	93.24	-1.47	10.05
211	Midday Sampling %	91.75	-1.44	1.89	96.80	-1.88	2.01	70.10	-1.75	2.11
Castledockerell 2011–12										
Ballycanew 2011–12										
Timoleague 2011–12										
Castledockerell 2012–13										
Ballycanew 2012–13										
Timoleague 2012–13										

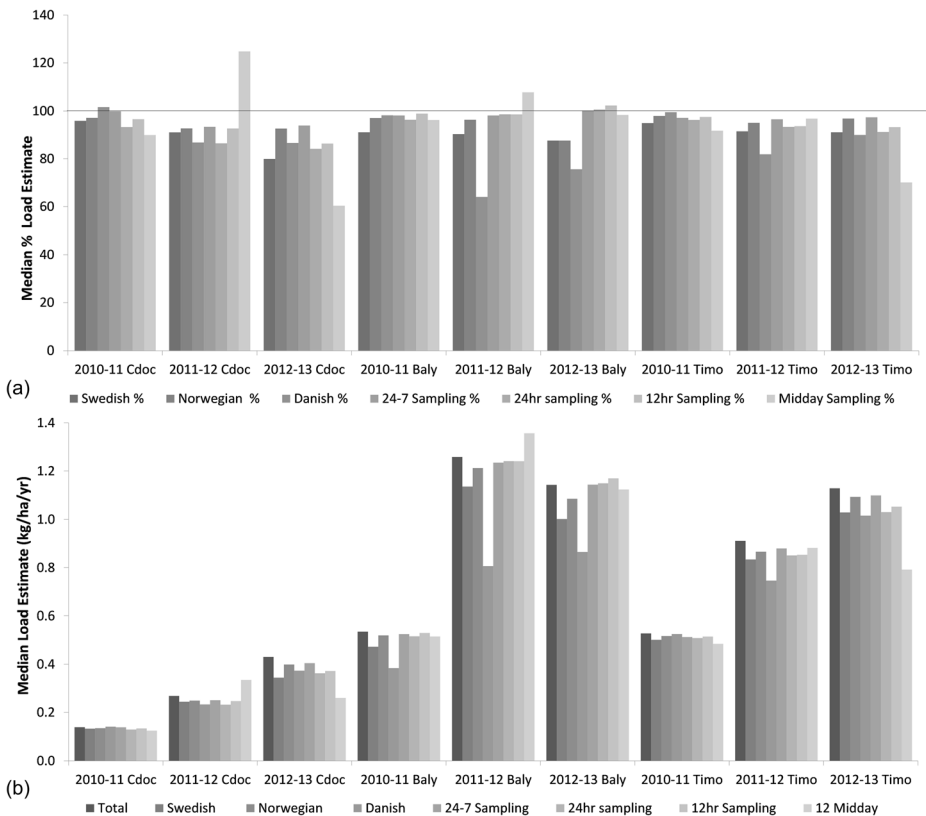


Fig. 2 Summary results from the simulated sampling approaches expressed as median % (a) and load/ha/yr (b). Catchment abbreviations are Castledockerell (Cdoc), Ballycanew (Baly) and Timoleague (Timo)

catchments. However, the difference between the implementation of this approach here (where twice the discharge for 1 week in advance is used to ‘forecast’ flow in the 2 week period) and the 2-week weather forecast used in practice should be considered. Though greatly improved since the late 1990s, when satellite data for northern Europe became available, rainfall forecasts still achieve only 70% accuracy over a 7 day period and about 40% for a 10 day forecast (Bauer et al. 2015). Local variations due to orographic controls should also be considered in implementing this approach and local meteorological records used to adapt sampling for each catchment.

The Swedish approach estimated 91.1, 90.3 and 87.6% of the load in Ballycanew, 95.9, 91.1 and 79.9% of the load in Castledockerell and 94.9, 91.5 and 91.1% of the total for Timoleague, over the years 2010–11, 2011–12 and 2012–13, respectively. Among the catchments Castledockerell had the lowest P export (154.3, 299.4 and 480.4 kg over the years 2010–11, 2011–12 and 2012–13, respectively) so failure to capture a small event may have a significant impact on the % of the load estimated compared to those catchments with higher exports (Ballycanew). For all composite sampling approaches the degree of variation among sample sets (Table 3) was low (maximum interquartile range = −0.63% to 1.06% for Norwegian and Swedish approaches; −7.8% to 11.44% for the Danish approach) due to the integrative nature of the composite abstraction of sub-samples being independent of the

starting time and the sampling frequencies capturing short-term storm transfers and diurnal variations.

3.2 Discrete Sampling Performance

Of the four discrete sampling approaches evaluated the 24–7 approach was most accurate overall, estimating 98.1, 98.1 and 100.1% of the load in Ballycanew, 99.9, 93.4 and 93.9% of the load in Castledockerell and 97.1, 96.6 and 97.4% of the load in Timoleague, over the years 2010–11, 2011–12 and 2012–13 respectively. Previous work has also shown how this approach preserves important process information during storm and non-storm periods (Jordan and Cassidy 2011). Load estimates for Ballycanew were the most accurate when expressed as a percentage but this was influenced by the much higher loads in that catchment, where each percentage of the annual load in Ballycanew equates to a much larger flux than for the other catchments (Fig. 2).

The variation in load estimates across sampling sets was greater than for composite sampling approaches due to the greater dependence on sampling initiation time where a maximum of 24 samples are taken over the course of a week and concentrations change rapidly through time. As such the interquartile ranges are higher than for the FWCS; between –14.2% and 20.6% of the median across the 24–7, 24 h, midday and 12 h sampling approaches. The 24–7 sampling approach is consistently the most accurate of the discrete approaches; most probably because the 7 h sampling interval ensures all times of day are covered over the course of a week and most storm events are captured. Midday sampling leads to over estimation of the load in both Castledockerell and Ballycanew in 2011–12 (median = 124% and 107.8% of loads, respectively) and a marked underestimation in Castledockerell and Timoleague in 2012–13 (median = 60.4% and 70.1%, respectively), which may be linked to persistent point sources of contamination that are likely more active at other times of day and therefore are not captured.

4 Conclusions

For high resolution capture of contaminant load in rivers, bankside analysers currently represent the best approach to monitoring, though not without limits in terms of the chemical fractions which can be analysed and instrument resolution. The considerable cost of resources, set-up, maintenance and data management have so far precluded implementation in a statutory monitoring context. As such, discrete or composite water sampling and laboratory analysis remain viable options for increasing accuracy at lower cost.

In this study proxy discrete sampling approaches for TP, particularly at daily and sub-daily frequencies, are shown to be accurate in rivers with a wide variation in base-flow index, though with some dependency on the starting time of the sampling programme. The cost of transport and laboratory analysis of samples is still likely to be high at this frequency if implemented over many catchments due to the number of samples likely required (>1200 per year for the 24/7 discrete approach but with 93–100% median load estimates). By comparison, proxy FWCS approaches appeared to deliver accuracy and with likely less resource required through analysis of a single weekly or fortnightly sample, which is representative of the flow integrated concentration profile across the period (26–52 samples per year). The success of the approach, however, does depend on the quality and accuracy of the autosampling unit (such

that small sample volumes can be consistently extracted) and the optimisation of the flow constant for each catchment and for temporal variations in rainfall and discharge.

The FWCS approaches evaluated here perform well, and for conservative chemical species, have application in situations where near-continuous monitoring is not feasible either due to costs or instrumentation. With the data and methods used, the Norwegian approach appeared to provide the better FWCS load estimates (87–98% median load estimates) – albeit using a proxy of this method and in reality using dynamic, forecasted flow constants. Utilisation of static FWCS and TWCS (for low flow) constants from extant hydrometric data showed that the Swedish approach performed almost as well (80–96% median load estimates). The trade-off between discrete and composite sampling approaches and costs of implementation could be assessed against current national statutory monitoring costs that compromise time for space and, with low cover of sample resolution across the flow range, will inevitably under-predict annual loads. Combinations of both FWCS and sub-daily discrete sampling in at least sentinel catchments (depending on what information is required) would be a development to consider for both policy and model assessments.

Acknowledgements We acknowledge data contributions from farmers and landowners, scientists, technicians and technologists affiliated to the Teagasc Agricultural Catchments Programme in Ireland. We thank technical staff from Norway, Denmark and Sweden for contributions during the development of this research.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Andersen HE, Kronvang B, Larsen SE (2005) Development, validation and application of Danish empirical phosphorus models nutrient mobility within river basins: a European perspective 304:355–365. <https://doi.org/10.1016/j.jhydrol.2004.07.039>
- Audet J, Martinsen L, Hasler B, de Jonge H, Karydi E, Ovesen NB, Kronvang B (2014) Comparison of sampling methodologies for nutrient monitoring in streams: uncertainties, costs and implications for mitigation. *Hydrol Earth Syst Sci* 18:4721–4731. <https://doi.org/10.5194/hess-18-4721-2014>
- Bauer P, Thorpe A, Brunet G (2015) The quiet revolution of numerical weather prediction. *Nature* 525:47–55. <https://doi.org/10.1038/nature14956>
- Bechmann M, Øgaard AF (2013) Water quality changes following intensive focus on mitigation methods to reduce phosphorus losses in the catchment of lake Vansjø, Norway. In: Sisák I (ed) *Proceedings of International Conference on Realistic Expectations for Improving European Waters.*, 12–14 October 2011, Keszthely, Hungary, 2013. ATON-Agrokémia és Talajtan ON-line [ATON-Agrochemistry and Soil Science ON-line: <http://www.aton.hu/documents/10156/a6acf379-62fd-4920-9b1c-847ed6393a38>, pp 103–117
- Bechmann M, Deelstra J, Stålnacke P, Eggstad HO, Øygarden L, Pengerud A (2008) Monitoring catchment scale agricultural pollution in Norway: policy instruments, implementation of mitigation methods and trends in nutrient and sediment losses. *Environ Sci Pol* 11:102–114. <https://doi.org/10.1016/j.envsci.2007.10.005>
- Brewin PA, Reynolds B, Stevens PA, Gee AS, Ormerod SJ (1996) The effect of sampling frequency on chemical parameters in acid-sensitive streams. *Environ Pollut* 93:147–157. [https://doi.org/10.1016/0269-7491\(96\)00028-0](https://doi.org/10.1016/0269-7491(96)00028-0)
- Cassidy R, Jordan P (2011) Limitations of instantaneous water quality sampling in surface-water catchments: comparison with near-continuous phosphorus time-series data. *J Hydrol* 405:182–193. <https://doi.org/10.1016/j.jhydrol.2011.05.020>
- Deelstra J, Stenrød M, Bechmann M, Eggstad HO (2013) Discharge measurement and water sampling. In: Bechmann M, Deelstra J (eds) *Agriculture and environment-long term monitoring in Norway*. Akademika Publishing, Trondheim, pp 83–104

- Eisenreich SJ, Bannerman RT, Armstrong DE (1975) A simplified phosphorus analysis technique. *Environ Lett* 9:43–53. <https://doi.org/10.1080/00139307509437455>
- Fealy RM et al (2010) The Irish agricultural catchments programme: catchment selection using spatial multi-criteria decision analysis. *Soil Use Manage* 26:225–236. <https://doi.org/10.1111/j.1475-2743.2010.00291.x>
- Gustard A, Bullock A, Dixon JM (1992) Low flow estimation in the United Kingdom. Institute of Hydrology, Wallingford
- Halliday SJ, Wade AJ, Skeffington RA, Neal C, Reynolds B, Rowland P, Neal M, Norris D (2012) An analysis of long-term trends, seasonality and short-term dynamics in water quality data from Plynlimon, Wales. *Sci Total Environ* 434:186–200. <https://doi.org/10.1016/j.scitotenv.2011.10.052>
- Harris G, Heathwaite AL (2005) Inadmissible evidence: knowledge and prediction in land and riverscapes. *J Hydrol* 304:3–19. <https://doi.org/10.1016/j.jhydrol.2004.07.020>
- Jarvie HP, Sharpley AN, Scott JT, Haggard BE, Bowes MJ, Massey LB (2012) Within-river phosphorus retention: accounting for a missing piece in the watershed phosphorus puzzle. *Environ Sci Technol* 46:13284–13292. <https://doi.org/10.1021/es303562y>
- Jordan P, Cassidy R (2011) Technical note: assessing a 24/7 solution for monitoring water quality loads in small river catchments. *Hydrol Earth Syst Sci* 15:3093–3100. <https://doi.org/10.5194/hess-15-3093-2011>
- Kronvang B, Bruhn AJ (1996) Choice of sampling strategy and estimation method for calculating nitrogen and phosphorus transport in small lowland streams. *Hydrol Process* 10:1483–1501. [https://doi.org/10.1002/\(SICI\)1099-1085\(199611\)10:11<1483::AID-HYP386>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1099-1085(199611)10:11<1483::AID-HYP386>3.0.CO;2-Y)
- Kyllmar K (2009) Transporter av kväve och fosfor i vattendrag – inverkan av metodik vid vattenprovtagning. Jämförelse av vattenanalyser från manuell respektive flödesproportionell vattenprovtagning i åtta Intensivtypområden. Swedish University of Agricultural Sciences, Department of Soil and Environment, Uppsala
- Kyllmar K, Forsberg LS, Andersson S, Mårtensson K (2014) Small agricultural monitoring catchments in Sweden representing environmental impact agriculture. *Ecosyst Environ* 198:25–35. <https://doi.org/10.1016/j.agee.2014.05.016>
- Littlewood IG, Watts CD, Custance JM (1998) Systematic application of United Kingdom river flow and quality databases for estimating annual river mass loads (1975–1994). *Sci Total Environ* 210–211:21–40. [https://doi.org/10.1016/S0048-9697\(98\)00042-4](https://doi.org/10.1016/S0048-9697(98)00042-4)
- Macintosh KA, Jordan P, Cassidy R, Arnscheidt J, Ward C (2011) Low flow water quality in rivers; septic tank systems and high-resolution phosphorus signals. *Sci Total Environ* 412–413:58–65. <https://doi.org/10.1016/j.scitotenv.2011.10.012>
- Marcé R, Schiller D, Aguilera R, Martí E, Bernal S (2018) Contribution of hydrologic opportunity and biogeochemical reactivity to the variability of nutrient retention in river networks. *Glob Biogeochem Cycles*. <https://doi.org/10.1002/2017GB005677>
- Melland AR, Fenton O, Jordan P (2018) Effects of agricultural land management changes on surface water quality: a review of meso-scale catchment research. *Environ Sci Pol* 84:19–25. <https://doi.org/10.1016/j.envsci.2018.02.011>
- Mellander P-E, Melland AR, Jordan P, Wall DP, Murphy PNC, Shortle G (2012) Quantifying nutrient transfer pathways in agricultural catchments using high temporal resolution data. *Environ Sci Pol* 24:44–57. <https://doi.org/10.1016/j.envsci.2012.06.004>
- Minaudo C et al (2017) Nonlinear empirical modeling to estimate phosphorus exports using continuous records of turbidity and discharge. *Water Resour Res* 53:7590–7606. <https://doi.org/10.1002/2017WR020590>
- Murphy PNC, Mellander PE, Melland AR, Buckley C, Shore M, Shortle G, Wall DP, Treacy M, Shine O, Mechan S, Jordan P (2015) Variable response to phosphorus mitigation measures across the nutrient transfer continuum in a dairy grassland catchment. *Agric Ecosyst Environ* 207:192–202. <https://doi.org/10.1016/j.agee.2015.04.008>
- Neal C, Reynolds B, Rowland P, Norris D, Kirchner JW, Neal M, Sleep D, Lawlor A, Woods C, Thacker S, Guyatt H, Vincent C, Hockenhull K, Wickham H, Harman S, Armstrong L (2012) High-frequency water quality time series in precipitation and streamflow: from fragmentary signals to scientific challenge. *Sci Total Environ* 434:3–12. <https://doi.org/10.1016/j.scitotenv.2011.10.072>
- Nimick DA, Gammons CH, Parker SR (2011) Diel biogeochemical processes and their effect on the aqueous chemistry of streams: A review. *Chem Geol* 283:3–17. <https://doi.org/10.1016/j.chemgeo.2010.08.017>
- Pellerin BA et al (2016) Emerging tools for continuous nutrient monitoring networks: sensors advancing science and water resources protection. *JAWRA J Am Water Resour Assoc*. <https://doi.org/10.1111/1752-1688.12386>
- Quilbé R, Rousseau AN, Duchemin M, Poulin A, Gangbazo G, Villeneuve J-P (2006) Selecting a calculation method to estimate sediment and nutrient loads in streams: application to the Beauvillage River (Québec, Canada). *J Hydrol* 326:295–310. <https://doi.org/10.1016/j.jhydrol.2005.11.008>

- Rode M, Suhr U (2007) Uncertainties in selected river water quality data. *Hydrol Earth Syst Sci* 11:863–874. <https://doi.org/10.5194/hess-11-863-2007>
- Skeffington RA, Halliday SJ, Wade AJ, Bowes MJ, Loewenthal M (2015) Using high-frequency water quality data to assess sampling strategies for the EU water framework directive. *Hydrol Earth Syst Sci* 19:2491–2504. <https://doi.org/10.5194/hess-19-2491-2015>
- Stamm C, Jarvie HP, Scott T (2014) What's more important for managing phosphorus: loads, concentrations or both? *Environ Sci Technol* 48:23–24. <https://doi.org/10.1021/es405148c>
- Stevens RJ, Smith RV (1978) A comparison of discrete and intensive sampling for measuring the loads of nitrogen and phosphorus in the river main, County Antrim. *Water Res* 12:823–830. [https://doi.org/10.1016/0043-1354\(78\)90033-7](https://doi.org/10.1016/0043-1354(78)90033-7)
- Vervloet LSC, Binning PJ, Børgesen CD, Højberg AL (2018) Delay in catchment nitrogen load to streams following restrictions on fertilizer application. *Sci Total Environ* 627:1154–1166. <https://doi.org/10.1016/j.scitotenv.2018.01.255>
- Wall DP et al (2012) Evaluating nutrient source regulations at different scales in five agricultural catchments. *Environ Sci Pol* 24:34–43. <https://doi.org/10.1016/j.envsci.2012.06.007>
- Worrall F et al (2012a) The flux of DOC from the UK – predicting the role of soils, land use and net watershed losses. *J Hydrol* 448–449:149–160. <https://doi.org/10.1016/j.jhydrol.2012.04.053>
- Worrall F, Davies H, Burt T, Howden NJK, Whelan MJ, Bhogal A, Lilly A (2012b) The flux of dissolved nitrogen from the UK — evaluating the role of soils and land use. *Sci Total Environ* 434:90–100. <https://doi.org/10.1016/j.scitotenv.2012.01.035>